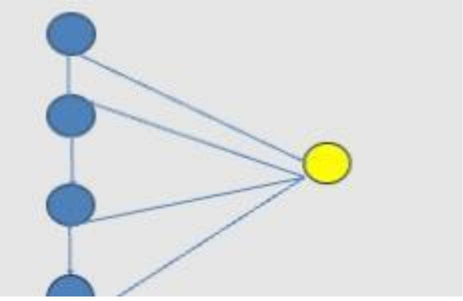
These days, People use Depp Neural Network(DNN) for classifying something or determining automatically. But many documents does not explain algorithms or structure of it, which make it hard to understand. However, we can alter DNN by other model like logistic regression or else if we well assume somethings. This kinds of modelling is useful for not only easy implementation but also fast calculation. So, I'll explain what is neural network like DNN or else and how we can alter it in some cases.

1.Neural Network(NN)



Many documents explain NN as a picture above, but this is just generalization of Factor Analysis(FA) which basis is non-linear function. FA is a kind of graphical modelling which assume there are some hidden common factors(yellow node) for explaining observed data. It is formulated as $X = WH + E$, which X is data vector, W is weight matrix(edges) H is a common factor vector and E is a error vector H can't explain. This require $EE^T = I$ for efficiency of W, H and if $E \sim N(0, I)$ then we can calculate W, H by matrix calculation. If the former does not holds, then we must use weighted metric. We can see a ith component x_i in X, then $x_i = \sum_k w_{ik} h_{kj} + e_i$.

The logic of NN is same as FA except H can be some non linear function like logistic function $\frac{e^{h_j}}{1+e^{h_j}}$ or else. Many people use NN for binary data to estimate what category of data should be allocated. But if we set group for data using some independent variables, then we can find that logistic regression is exactly same as NN. As I say in document about it, we can modify that function depending on our data. The logistic regression can be calculated by linear regression which is faster than backpropagation.

We talk about that NN is just graphical model which basis(or activation functions) are non-linear. Then why we add some layers on model, which is called DNN? As I said above, the nodes in layer are just some kind of functions which do some actions and this means that we only need to add them if we think it works in some sense. For example, suppose we want to classify pictures into some categories and we want programs to do it like us. Think about how we have done it before. We focus on some parts of object and use them to descriminate it with other objects. In this case, we can set 2 layers, which one is collection of nodes for extracting some traits in pictures like Support Vector Machine(SVM) and the other is set of nodes for classification. You can also use 3 layers by adding some nodes for preprocessing of features which we already get in first layers.

We can say that other NN algorithms like above. For examples, Recurrent Neural Network(RNN) is generalization of Hidden Markov Model(HMM). In HMM, there are some hidden nodes which we can't observe directly but they have some markov chain and they emit some data with probability p, which data can be observed by us. RNN is generalization of it and admit other model like non-markov system for hidden states. So, if we know systems of time series then we can use it by applying kalman-filter or Vector AutoRegression(VAR) for fast calculations.

To summary, Neural Network is just generalized Factor Analysis which non-linear function is used in hidden nodes and we don't need to use it if we think that some other model we already know is enough. Also, adding layer means we do some actions on data like data pre-processing, feature extracting, some complex hidden models or else. Anyway, the most important thing is that we should set appropriate model for our goal.